

Speed exponents of random walks on groups

Gideon Amir

Bálint Virág

March 1, 2013

Abstract

For every $3/4 \leq \beta < 1$ we construct a finitely generated group so that the expected distance of the simple random walk from its starting point satisfies $\mathbf{E}|X_n| \asymp n^\beta$. In fact, the speed can be set precisely to equal any nice prescribed function up to a constant factor.

1 Introduction

The central question in the theory of random walks on groups is how symmetries of the underlying space give rise to structure and rigidity of the random walks. For example, for nilpotent groups, it follows easily from bounds in Hebisch and Saloff-Coste (1993) that random walks have diffusive behavior, namely that the rate of escape, defined as the expected distance of the walk from the identity satisfies

$$\mathbf{E}|X_n| \asymp n^{1/2}.$$

(See also Thompson (2010) for more classes with this behavior.) On nonamenable groups, on the other hand, we have

$$\mathbf{E}|X_n| \asymp n.$$

Starting with the works of Kesten (1959), the rate of escape has been connected to many other properties of the group, including the spectral radius, the volume growth, rate of entropy growth, and the Liouville property. More recently, Austin, Naor and Peres (2009) and Naor and Peres (2008) have established a beautiful connection between the rate of escape and embeddings of groups into Hilbert space, see (17).

The notation $f \asymp g$ denotes that there exist constants c, C independent of n such that $cf(n) \leq g \leq Cf(n)$.

Their work renewed interest in a long-standing question, attributed to Vershik, namely, what rates of escape are possible. In the formulation of Peres (personal communication):

For what β is there a random walk on a group with $\mathbf{E}|X_n| = n^{\beta+o(1)}$?

The only known answers to this question were given by Erschler (1999), who realized that for a random walk on a group G satisfying $\mathbf{E}|X_n| \asymp n^\beta$, the wreath product $\mathbb{Z} \wr G$ has $\mathbf{E}|X_n| \asymp n^{\frac{\beta+1}{2}}$.

Thus Erschler showed that the canonical random walk on the $k-1$ -times iterated wreath product of \mathbb{Z} satisfies

$$\mathbf{E}|X_n| \asymp n^{1-2^{-k}}.$$

The main goal of this paper is to show that in fact, all speed exponents between $3/4$ and 1 are achievable in residually finite groups. In fact, we show much more: our groups allow us to adjust precisely $\mathbf{E}|X_n| \asymp f(n)$ for a general class of functions f . For this, we use the permutation wreath product of \mathbb{Z} and the action S of the so called piecewise mother groups \mathcal{M}_m on the boundary of their tree; see Section 6 for definitions.

Theorem 1. *For any $\gamma \in [3/4, 1)$ and any function $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ satisfying $f(1) = 1$ and the log-Lipshitz condition that for all real $a, n \geq 1$*

$$a^{3/4}f(n) \leq f(an) \leq a^\gamma f(n)$$

there is a bounded sequence (m_ℓ) such the random walk on $\mathbb{Z} \wr_S \mathcal{M}_m$ satisfies that with constants depending on γ only, $\mathbf{E}|X_n| \asymp f(n)$.

For the theorem and throughout this paper, all random walks are symmetric and the step distribution has finite support. While the actual group construction for Theorem 1 is simple, the theory behind the analysis is more sophisticated and has many potential applications. It is based on the analysis of permutation wreath products (see Section 4 for definitions).

Theorem 2. *Let Y_n be a random walk on a group Γ which acts on a set S . Fix $o \in S$, and let p_n be the probability that the random walk $o.Y_n$ on S does not return to o in the first n steps. Assume that there is a constant c so that for all $n \geq 1$ we have*

$$\mathbf{E}|Y_n| + H[\{o, o.Y_1^{-1}, \dots, o.Y_n^{-1}\}] + (p_0 + \dots + p_n) \leq cnp_n, \quad (1)$$

*where the H term is the entropy of a random set called the **inverted orbit**. Let Λ be an infinite group and let $\lambda(n)$ be the expected displacement for a generating random walk in*

Λ . Then the switch-walk-switch random walk X_n on the permutation wreath product $\Lambda \wr_S \Gamma$ satisfies

$$\mathbf{E}|X_n| \asymp np_n \lambda(1/p_n).$$

The main reason this theorem is useful is that it gives control of the expected displacement $\mathbf{E}|X_n|$, usually a difficult quantity, in terms of the function p_n , which is often easier to compute. The condition (1) ensures that the action is significant enough to determine the speed of the switch-walk-switch random walk.

For a simple example, consider $\Gamma = \mathbb{Z}$, with the free (Cayley) action. Then we are back to the ordinary wreath product, $p_n \asymp n^{-1/2}$, and $\mathbf{E}|Y_n| \asymp n^{1/2}$, and the entropy in (1) is of order $\log n$. Repeated application of Theorem 2 then recovers the result of Erschler (1999).

Because of Theorem 2, understanding speed and entropy in wreath products can boil down to understanding return times of random walks on Schreier graphs, a much easier task. Finally, we make this easier by constructing a family of Schreier graphs on which the random walk simply projects to nonnegative birth-and-death Markov chains. These are chains on the nonnegative integers where only nearest-neighbor moves are allowed; the random walk there can be analyzed many ways, including electrical network theory. The corresponding action satisfies the condition (1), and Theorem 1 holds since p_n can be adjusted sufficiently precisely (in fact to any function as specified in Theorem 1 with $3/4$ replaced by $1/2$.) Then by Theorem 2, the $\mathbf{E}|X_n|$ can be determined up to a constant from the expected displacement in Λ for any infinite group Λ .

Our analysis is precise enough to give a quantitative improvement on Brioussell's theorem on entropy. In Brioussell (2011a) it is shown that entropy exponents between $1/2$ and 1 are realizable; the same analysis can be extended to $H(X_n) = f(n)n^{o(1)}$ for a general class of functions f . Here we show that in fact the random walk entropy can also be adjusted much more precisely, up to a constant factor.

Theorem 3. *For any $\gamma \in [1/2, 1)$ and any function $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ satisfying $f(1) = 1$ and the log-Lipshitz condition that for all real $a, n \geq 1$*

$$a^{1/2}f(n) \leq f(an) \leq a^\gamma f(n)$$

there is a bounded sequence (m_ℓ) such the random walk on $\mathbb{Z}_2 \wr_S \mathcal{M}_m$ satisfies $H(X_n) \asymp f(n)$, and the implied constants depend on γ only.

The piecewise mother groups \mathcal{M}_m that we are using are related to those in the paper of Brioussell; the main difference is that our analysis is based on a careful geometric under-

standing of the birth-and-death chain structure of the underlying Markov chains; also, it is key in our analysis of the speed to let Λ be an infinite group.

The groups \mathcal{M}_m are versions of automaton groups. The most famous of such groups is the celebrated group of Grigorchuk (1984). The rate of escape on automaton groups was first studied by Bartholdi and Virág (2005), who used its sublinearity to show that the Basilica group is amenable. These results were extended to so-called degree zero automaton groups by Bartholdi, Kaimanovich and Nekrashevych (2010), and for degree one groups by Amir, Angel and Virág (2009). In Amir and Virág (2011), it was shown that degree three groups typically have linear rate of escape.

The idea to use permutational wreath product over these groups is due to Bartholdi and Erschler (2010); they used the Grigorchuk group as a base group to determine the exponent of the stretched exponential growth of its permutational wreath product; to do so for any intermediate growth group had been an open question. Brieussel (2011b) used a piecewise version of this construction (the likeness of which was used earlier by Erschler (2006) to give examples for Follner functions) to show that all exponents within a range are possible. Later, Kassabov and Pak (2011) gave a different construction. Most recently, Brieussel (2011a) used the same groups we use to show that entropy exponents can freely vary between $1/2$ and 1 .

The permutational wreath product can be thought of as a relative of the usual wreath product $\Lambda \wr \Gamma$, also known as the lamplighter product of Λ and γ . Thus we borrow terminology and think of Λ as the "lamp group" and of the switch-walk-switch walk as a type of Lamplighter walk. This analogy is limited though and must be used with care. For example, perhaps surprisingly, the lamplighter corresponding to a permutation wreath product does not light lamps along the random walk path $o.Y_n$ on the corresponding Schreier graph with root o . Instead, it lights the lamps along the inverted orbit $o.Y_n^{-1}$, which may have a very different distribution (see Section 4).

The rest of the paper is built as follows: In Section 2, we analyze the distribution of the size of the inverted orbit. In Section 3 we give a precise version of the well-known inequality connecting random walk entropy and speed based on the bounds by Varopoulos and Carne. In Section 4, we define permutational wreath products and develop our bounds for speed and entropy in general permutational wreath products. In Section 5, we introduce and analyze the assembly line, a birth-and-death Markov chain which will be the basis of our construction; these are simpler to understand than the underlying groups. In Section 6 we introduce the underlying groups, prove specific estimates for the orbits in these groups and

then apply the bounds of the previous sections to deduce the three main theorems. Finally, we close the paper with some open questions.

2 Random walks and inverted orbits

Let $T_i, i \in \mathbb{Z}$ be a positive integer-valued i.i.d. sequence. The two sided-random walk with increments T_i has value $T_1 + \dots + T_i$ for $i \geq 1$ and $-T_{-1} - \dots - T_i$ for $i < 0$. Its image as i ranges over \mathbb{Z} is a random set called the two-sided renewal process. Its gap distribution is the law of T_i .

Proposition 4. *Let $Y_i = G_1 G_2 \dots G_i$ be a random walk on a group G , acting on a rooted set (S, o) . Let*

$$V_t = \{i \geq 0 : o.Y_i^{-1} = o.Y_t^{-1}\},$$

namely the times at which the inverted orbit visits the same point as at time t . Then V_t has the same distribution as $(t + \rho) \cap [0, \infty)$ where ρ is the two-sided renewal process whose gap distribution is the law of the first return time of the Markov chain $o.Y_n$.

Proof. Fix n . Let $R \subset [1, t]$ and $S \subset [1, n - t]$ be sets of integer times, and consider the event $V_t \cap [0, n] = (t - R) \cup (t + S)$. We can write this event as the intersection of the events that

$$\{1 \leq r \leq t : o.Y_{t-r}^{-1} = o.Y_t^{-1}\} = R$$

and

$$\{1 \leq s \leq n - t : o.Y_{t+s}^{-1} = o.Y_t^{-1}\} = S.$$

The equality in the first event can be written as

$$o.G_{t-r}^{-1} \dots G_1^{-1} = o.G_t^{-1} \dots G_1^{-1}.$$

Multiplying from the right by $G_1 \dots G_{t-r}$ we get

$$o = o.G_t^{-1} \dots G_{t-r}^{-1}.$$

Similarly, the equality in the second event is equivalent to

$$o = o.G_{t+1} \dots G_{t+s}$$

Clearly, these specify the return times of two independent Markov chains. Such return times are renewal processes, and the gap distribution is given by the time of the first return. To

verify our claim, all we have to check is that the first return time T for the walk with the inverse steps has the same distribution as that for the random walk with normal steps. Note that $T = k$ is equivalent to

$$o.G_1^{-1} \cdots G_k^{-1} = o, \quad (2)$$

and

$$o.G_1^{-1} \cdots G_\ell^{-1} \neq o, \text{ for all } \ell < k.$$

If the first equality holds, then the second one is equivalent to

$$o.G_{\ell+1}^{-1} \cdots G_k^{-1} \neq o, \text{ for all } \ell < k$$

or, multiplying over, to

$$o \neq o.G_k \cdots G_{\ell+1}, \text{ for all } \ell < k.$$

Now, together with (2), this specifies the first return time for the walk with group increments in the reverse order: G_k, \dots, G_1 . \square

Let Q_n be the (random) occupation measure of the inverted orbit of the first n steps. More precisely, for $s \in S$ we have $Q_n(s) = \sum_{i=0}^n \mathbf{1}(o.Y_i^{-1} = s)$, the number of visits of the inverted orbit to s . We will often think of n as fixed and call Q_n simply the inverted orbit. The support of Q_n is will be called the range. Let $|Q_n|$ denote the size of the range.

Corollary 5. *Let T denote the first return time for the random walk $o.Y_n$. Then we have*

$$\mathbf{E}|Q_n| = \sum_{i=0}^n \mathbf{P}(T > i) = \mathbf{E}[T \wedge (n+1)].$$

Proof. Let B_i the event that time i is the last visit to $o.Y_i$ up to time n . Note that by Proposition 4 $\mathbf{P}(B_i) = \mathbf{P}(T > n - i)$. We have

$$\mathbf{E}|Q_n| = \mathbf{E} \sum_{i=0}^n \mathbf{1}_{B_i} = \sum_{i=0}^n \mathbf{P}(B_i)$$

Changing variables with $j = n - i$ we get

$$\mathbf{E}|Q_n| = \sum_{j=0}^n \mathbf{P}(T > j) = \sum_{j=0}^{\infty} \mathbf{P}(T \wedge (n+1) > j)$$

so the last equality follows from the tail sum formula for expectation. \square

Corollary 6. *Let T denote the first return time for the random walk $o.Y_n$. Then for $0 \leq t \leq n$ we have*

$$\mathbf{P}(Q_n(o.Y_t^{-1}) \leq k) \geq 1 - 2\mathbf{P}(T \leq n)^{k/2}.$$

Proof. By Proposition 4, we have

$$\mathbf{P}(Q_n(o.Y_t^{-1}) \leq k) = \mathbf{P}(A + B < k)$$

where A, B are the number of renewals up to and including time t and $n - t$, respectively in two independent renewal processes with gap distribution given by the law of T . This gives the lower bound

$$\mathbf{P}(Q_n(o.X_t) \leq k) \geq 1 - \mathbf{P}(A \geq k/2) - \mathbf{P}(B \geq k/2)$$

If $A \geq k/2$, then the first $k/2$ gaps are all of size at most $t \leq n$. So, by the independence of gaps, we can bound

$$\mathbf{P}(A \geq k/2) \leq \mathbf{P}(T \leq n)^{k/2}.$$

The same argument for B now concludes the proof. \square

3 Entropy and speed

Recall that the **entropy** of a finite non-negative measure μ supported on a countable set G is defined by

$$H(\mu) = \sum_{x \in G} -\mu(x) \log \mu(x),$$

where by convention $0 \log 0 = 0$. The entropy $H(X)$ of a discrete random variable X is given by the entropy of its distribution; the entropy $H(X_1, \dots, X_n)$ of several random variables is given by the entropy of the joint distribution of the X_i . In order to define conditional entropy for two random variables X, Y , let $f(y)$ denote the entropy of the conditional distribution of X given $Y = y$. Then the **conditional entropy of X given Y** is defined as $H(X|Y) := \mathbb{E}f(Y)$.

The conditional entropy satisfies

$$H(X, Y) = H(X|Y) + H(Y).$$

A useful and easy fact is that among measures supported on a given finite set, the one having maximal entropy is the uniform measure on that set.

Let $h(n)$ be the entropy of the n -step random walk in the group Λ . Recall that h is concave (Kaimanovich and Vershik (1983), Proposition 1.3) and hence by linear approximation we can extend it to a concave function on \mathbb{R}_+ .

Lemma 7. *For every $n \geq 1$ the function $x \mapsto xh(n/x)$ is concave.*

Proof. By the concavity of entropy, we have

$$\frac{x}{x+y}h(n/x) + \frac{y}{x+y}h(n/y) \leq h\left(\frac{x}{x+y}n/x + \frac{y}{x+y}n/y\right) = h\left(\frac{2n}{x+y}\right)$$

equivalently,

$$\frac{xh(n/x) + yh(n/y)}{2} \leq \frac{x+y}{2}h\left(\frac{2n}{x+y}\right)$$

which is the claimed concavity. \square

The following is a precise version of the well-known relationship between entropy and expected distance (Erschler (2003)) that follows from the bounds of Varopoulos (1985) and Carne (1985).

Proposition 8. *Consider a random walk X_n on a weighted graph started at a vertex o , and let $|\cdot|$ denote graph distance from o . Then we have*

$$\mathbf{E}|X_n| \leq 4\sqrt{n\left(H(X_n) \vee \max_v \frac{\deg v}{\deg o}\right)}.$$

Proof. Let $r \geq 0$ an integer. By leaving out nonnegative terms from the definition, we write

$$H(X_n) \geq - \sum_{v \in V, |v| \geq r} \mathbf{P}(X_n = v) \log \mathbf{P}(X_n = v). \quad (3)$$

By the bounds of Varopoulos (1985) and Carne (1985), with $\eta = \max_v \frac{\deg v}{\deg o}$ and $\kappa = \log(2\sqrt{\eta})$ we have

$$\log \mathbf{P}(X_n = v) \leq \kappa - |v|^2/2n.$$

Substituting this into (3) and then summing over v with $|v| \geq r$ we get

$$H(X_n) \geq \mathbf{P}(|X_n| \geq r) \left(\frac{r^2}{2n} - \kappa \right).$$

For integer $r > \sqrt{2\kappa n}$ this gives

$$\mathbf{P}(|X_n| \geq r) \leq \frac{2nH(X_n)}{r^2 - 2\kappa n} := q(r).$$

Since $q(r)$ is decreasing, this also holds for all real $r > \sqrt{2\kappa n}$. Indeed,

$$\mathbf{P}(|X_n| \geq r) = \mathbf{P}(|X_n| \geq \lceil r \rceil) \leq q(\lceil r \rceil) \leq q(r).$$

We now write

$$\mathbf{E}|X_n| = \int_0^\infty \mathbf{P}(|X_n| \geq r) dr \leq r_0 + \int_{r_0}^\infty \frac{2nH(X_n)}{r^2 - 2\kappa n} dr$$

and substitute $r_0 = 2\sqrt{nH(X_n) \vee n\kappa}$, so that the numerator of the integrand is at least $r^2/2$. This gives the upper bound

$$r_0 + \int_{r_0}^{\infty} \frac{4nH(X_n)}{r^2} \leq 4\sqrt{nH(X_n) \vee n\kappa}$$

Since $\eta \geq 1$ and so $\kappa = \log(2\sqrt{\eta}) < \eta$ the claim follows. \square

4 Speed in permutation wreath products

We consider the following setup. Let Γ be a finitely generated countable infinite group acting on a set S . We single out an element $o \in S$ and call it the root. Let μ be a finitely supported symmetric measure on Γ . Let Λ be a finitely generated countable (possibly finite) group. The **permutation wreath product** $\Lambda \wr_S \Gamma$ is the semidirect product of Λ^S with Γ acting on it by permuting the coordinates. The multiplication rule, for $\ell, \ell' \in (\Lambda^S)$ and $g, g' \in \Gamma$ is

$$(\ell, g)(\ell', g') = (\ell\ell'^{g^{-1}}, gg')$$

where $\ell'^{g^{-1}}$ is defined by $\ell'^{g^{-1}}(s) = \ell'(s.g)$.

A **switch** is a random element of $\Lambda \wr_S \Gamma$ of the form (\bar{L}, id_Γ) , where

$$\bar{L}(s) = \begin{cases} id_\Lambda & \text{if } s \neq o, \\ L & \text{if } s = o, \end{cases}$$

where L is a random element of Λ chosen from a fixed symmetric finitely-supported measure.

We consider the random walk

$$X_n = \prod_{i=1}^n \bar{L}_i G_i \bar{L}'_i$$

(called the **switch-walk-switch** random walk) on the permutation wreath product. Here the G_i are independent choices from the measure on Γ , and the L_i, L'_i are independent choices from the measure on Λ . We have $X_i = (\cdot, Y_i)$ where $Y_i = G_1 \cdots G_i$.

Keeping with the tradition for ordinary wreath products, we call the group $\Lambda \wr_S \Gamma$ the Lamplighter group, and the above walk the lamplighter walk.

Unlike the usual lamplighter group, where there is a walker on the (Cayley) graph who keeps switching lights along its path, this walk cannot be interpreted as a lamplighter moving along the Schreier graph of the action G on S . This is because the switches happen at locations $o, o.G_1^{-1}, o.G_2^{-1}G_1^{-1}, o.G_3^{-1}G_2^{-1}G_1^{-1}$, a sequence that is not necessarily Markovian

(or even connected in the Schreier graph). Note that $o, o.G_1^{-1}, o.G_1^{-1}G_2^{-1}, o.G_1^{-1}G_2^{-1}G_3^{-1}$ would be a Markovian random walk on the Schreier graph, but our case is different. Following Bartholdi and Erschler (2010) we call the sequence $o, o.G_1^{-1}, o.G_2^{-1}G_1^{-1}, o.G_3^{-1}G_2^{-1}G_1^{-1}, \dots$ the inverted orbit of o under the walk Y_n (or alternatively of the walk X_n as they have the same action on S .) In particular, we get the useful observation, which will be exploited later, that the inverted orbit of X_n depends only on Y_n and not on the lamp group or the switch steps chosen.

A symmetric finitely-supported measure on a group defines a Cayley graph. Let $|g|$ denote the corresponding graph distance of g from the identity.

With L_i random walk increments on Λ as above, define

$$\underline{\lambda}(t) = \inf_{n \geq t} \mathbf{E}|L_1 \cdots L_n|, \quad \bar{\lambda}(t) = \max_{n \leq t} \mathbf{E}|L_1 \cdots L_n|.$$

A simple computation checks that both $\underline{\lambda}$ and $\bar{\lambda}$ are subadditive: $\lambda(t+s) \leq \lambda(t) + \lambda(s)$. Also, unlike the ordinary expected distance, both $\underline{\lambda}$ and $\bar{\lambda}$ are nondecreasing (by definition). Lemma 4.1 in Lee and Peres (2009) implies that

$$\frac{1}{2}(\bar{\lambda}(n) - 1) \leq \mathbf{E}|L_1 \cdots L_n| \leq 2\underline{\lambda}(n) - 1 \quad (4)$$

The following theorem connects speed on Λ to that on the permutation wreath product. The main feature is that it only uses the tail of the first return time T of the random walk on the Schreier graph of (S, Γ) . We emphasize that T relates to the ordinary random walk.

Theorem 9. *The random walk X_n on the permutation wreath product satisfies for $n \geq 1$*

$$\mathbf{E}|X_n| \geq \frac{np\underline{\lambda}(1/p)}{16}.$$

for every $p \leq \mathbf{P}(T > n)$.

Proof. Note that the number of switches of the light at s by time n given by

$$\tilde{Q}_n(s) = 2Q_n(s) - \mathbf{1}_{\{o\}}(s) - \mathbf{1}_{\{o.Y_n^{-1}\}}(s) \geq Q_n(s) \quad (5)$$

for $n \geq 1$. Since all these lights have to be turned off to return to the identity, we have

$$\mathbf{E}|X_n| \geq \mathbf{E} \sum_{s \in S} \underline{\lambda}(Q_n(s))$$

Let k be a constant whose value will be decided later on. Call an element $s \in S$ thin if $Q_n(s) \leq k$, and call a time $0 \leq t \leq n$ thin if $o.Y_t^{-1}$ is thin. We can bound the right hand side below as

$$\mathbf{E} \sum_{s \in S, s \text{ thin}} \underline{\lambda}(Q_n(s)).$$

Let N be the number of thin times. We want to bound the quantity in the expectation in terms of N . This gives the following minimization problem. Minimize $\sum \underline{\lambda}(x_i)$ given that $\sum x_i = N$ and the $x_i < k$ are nonnegative integers. Note that since $\underline{\lambda}$ is subadditive, we can decrease the quantity in question without violating the constraints if we replace two x_i 's that are less than $k/2$ by their sum and zero. After repeating this procedure, we get that all but perhaps one of the nonzero x_i are greater or equal than $k/2$. Let j be the number of x_i at least $k/2$. Then we have $jk + k/2 \geq N$, so $j \geq N/k - 1/2 \geq N/(2k)$. This gives

$$\sum \underline{\lambda}(x_i) \geq \frac{N}{2k} \underline{\lambda}(k/2)$$

substituting this into the above, we get

$$\mathbf{E}|X_n| \geq \mathbf{E} \frac{N}{2k} \underline{\lambda}(k/2) = \frac{\mathbf{E}N}{2k} \underline{\lambda}(k/2) \quad (6)$$

it remains to compute

$$\mathbf{E}N = \sum_{t=0}^n \mathbf{P}(Q_n(o, Y_t^{-1}) \leq k) \geq n(1 - 2(1 - \mathbf{P}(T > n))^{k/2}) \geq n(1 - 2(1 - p)^{k/2})$$

by Corollary 6. Setting $k = 2/p$ gives the lower bound of $n/4$, and substituting into (6) we get

$$\mathbf{E}|X_n| \geq \frac{np \underline{\lambda}(1/p)}{16}$$

as claimed. \square

Analogous to Theorem 9 about speed, we have an upper bound for entropy in the wreath product. Again, the relation involves the return time of the random walk in the Schreier graph, since by Corollary 5

$$\mathbf{E}|Q_n| = \sum_{i=0}^n \mathbf{P}(T > i) = \mathbf{E}[T \wedge (n+1)].$$

Theorem 10. *We have*

$$H(X_n) \leq H(Y_n) + H(\text{supp } Q_n | |Q_n|) + 2\mathbf{E}|Q_n| \left(h\left(\frac{n}{\mathbf{E}|Q_n|}\right) + \log(n+1) \right).$$

and the expression on the right is increasing as a function of $\mathbf{E}|Q_n|$.

Proof. Write $X_n = (\mathcal{L}_n, Y_n)$, a pair containing a lamp configuration and a walk Y_n on Γ .

$$H(X_n) \leq H(Y_n) + H(\mathcal{L}_n) \leq H(Y_n) + H(\mathcal{L}_n | Q_n) + H(Q_n) \quad (7)$$

We first consider the entropy of Q_n .

$$\begin{aligned} H(Q_n) &= H(\text{supp } Q_n) + H(Q_n | \text{supp } Q_n) \\ &\leq H(\text{supp } Q_n | |Q_n|) + H(|Q_n|) + \mathbf{E}|Q_n| \log(n+1) \end{aligned} \quad (8)$$

the second inequality follows since $1 \leq Q_n(s) \leq n+1$ for every $s \in \text{supp } Q_n$. Further, the random variable $|Q_n|$ is between 1 and $n+1$, so its entropy is at most $\log(n+1)$.

We now turn to $H(\mathcal{L}_n | Q_n)$. Note that given Q_n , each entry of $\mathcal{L}_n(s)$ is distributed as independent samples from a random walk on Λ of length $\tilde{Q}_n(s) \leq 2Q_n(s)$ (see (5)), which is at most

$$H(\mathcal{L}_n | Q_n) \leq \mathbf{E} \sum_s h(\tilde{Q}_n(s))$$

now we use the monotonicity, concavity, and subadditivity of h (Kaimanovich and Vershik (1983) propositions 1.1 and 1.3) to get

$$\sum_s h(\tilde{Q}_n(s)) \leq \sum_s 2h(Q_n(s)) \leq 2|Q_n|h(n/|Q_n|)$$

By Lemma 7, we have that $x \mapsto xh(n/x)$ is concave, so by Jensen's inequality

$$H(\mathcal{L}_n | Q_n) \leq 2\mathbf{E}|Q_n|h(n/\mathbf{E}|Q_n|).$$

We summarize

$$H(X_n) \leq H(Y_n) + 2\mathbf{E}|Q_n|h(n/\mathbf{E}|Q_n|) + H(\text{supp } Q_n | |Q_n|) + \mathbf{E}|Q_n| \log(n+1) + \log(n+1).$$

Using $|Q_n| \geq 1$ the claimed inequality follows. Since the function $x \mapsto xh(n/x)$ is concave and nonnegative, it must be increasing. \square

Remark 11.

1. Given the size $|Q_n|$ of the support of Q_n , the support can be described by a sequence of $|Q_n|$ elements from the ball $B_n(o)$ of radius n in the Schreier graph of (S, Γ) . This gives

$$H(\text{supp } Q_n | |Q_n|) \leq \mathbf{E}|Q_n| \log |B_n(o)|.$$

However, as we will see in Section 5, for the groups we will be interested in one can give a better bound, linear in $|Q_n|$.

2. If the lamp group Λ is finite, one may replace (7) by

$$H(X_n) \leq H(Y_n) + H(\mathcal{L}_n) \leq H(Y_n) + H(\mathcal{L}_n | \text{supp } Q_n) + H(\text{supp } Q_n)$$

which in turn gives

$$H(X_n) \leq H(Y_n) + \mathbf{E}|Q_n| \log |\Lambda| + H(\text{supp } Q_n) \quad (9)$$

as every word in Λ can be described by $\log |\Lambda|$ bits.

The following lemma estimates the size of the inverted orbit of a random walk on a Schreier graph in terms of the speed of random walk on the permutation wreath product with \mathbb{Z}_2 . Together with Theorem 13 it usually gives a better upper bound on the speed (of the random walk X_n on $\Lambda \wr_S G$) than applying the Varopolous-Carne bound of Proposition 8 directly to the entropy bound of Theorem 10.

Lemma 12. *Consider a random walk Y_n on a Schreier graph S of a group G with inverted orbit Q_n . Then Y_n can be realized as the position of the lighter in the switch-walk-switch random walk X'_n on the permutation wreath product $\mathbb{Z}_2 \wr_S G$. There exists a word in G that is equivalent to Y_n whose inverted orbit includes $\text{supp } Q_n$, and whose expected length is at most $3\mathbf{E}|X'_n|$.*

Proof. Note that given Q_n , each lamp in $\text{supp } Q_n$ is turned on or off with probability half each, independently.

If \mathcal{L}'_n is the lamp configuration at time n , then given Q_n , the distribution of \mathcal{L}'_n and the inverted configuration \mathcal{L}''_n defined as

$$\mathcal{L}''_n(s) = \mathbf{1}(Q_n(s) > 0) - \mathcal{L}'_n(s)$$

is the same. Let $X''_n = (\mathcal{L}''_n, Y_n)$. By the above, X''_n has the same distribution as X'_n .

Since in X'_n all the lamps in \mathcal{L}'_n are turned on, there is a word W' of length at most $|X'_n|$, equivalent to X'_n , whose inverted orbit includes $\text{supp } \mathcal{L}'_n$. Similarly, there is a word W'' of length at most $|X''_n|$ whose inverted orbit includes $\text{supp } \mathcal{L}''_n$. Then the inverted orbit of the word $W''W'^{-1}W'$ contains $\text{supp } Q_n$, it is equivalent to X'_n and its expected word length is at most

$$2\mathbf{E}|X''_n| + \mathbf{E}|X'_n| = 3\mathbf{E}|X'_n|. \quad \square$$

Theorem 13. *For every $q \geq \mathbf{E}|Q_n|$ we have*

$$\mathbf{E}|X_n| \leq 3\bar{\lambda}(n/q)q + 12\sqrt{n}\sqrt{H(Y_n) + H(\text{supp } Q_n) + \mathbf{E}|Q_n|}.$$

Proof. We have $|X_n| \leq A + B$, where A is the length of the shortest word equivalent to Y_n whose inverted orbit contains $\text{supp } Q_n$ and B is the number of lamp steps required to turn off the lamps.

We first bound B . Given Q_n , the expected number of lamp moves required to turn off the lamps is bounded above by

$$\mathbf{E}[B|Q_n] \leq \sum_{v \in \text{supp } Q_n} \bar{\lambda}(Q_n(v))$$

first we use the monotonicity of $\bar{\lambda}$ to round up each argument to the next integer multiple of $k = n/q$. The total sum of the new arguments is at most the total sum of the originals ones (i.e. $2n$) plus $|Q_n|k$. Then we use subadditivity to bound the sum by

$$\bar{\lambda}(k) \frac{2n + |Q_n|k}{k} = \bar{\lambda}(k)(2q + |Q_n|)$$

Taking expectations we get

$$\mathbf{E}B \leq \bar{\lambda}(n/q) 3q.$$

Regarding A , Lemma 12 gives us

$$\mathbf{E}A \leq 3\mathbf{E}|X'_n|,$$

where X'_n is the switch-walk-switch random walk on $\mathbb{Z}_2 \wr_S G$. For this, we use the entropy bound (9) and the Varopoulos-Carne bounds in Proposition 8 to get

$$\mathbf{E}|X'_n| \leq 4\sqrt{n}(H(Y_n) + H(\text{supp } Q_n) + (\log 2)\mathbf{E}|Q_n|)^{1/2}$$

This completes the proof. □

We summarize the results of this section and deduce some conclusions using the Varopoulos-Carne bounds in Proposition 8. The second upper bound on the speed is usually sharper.

Corollary 14. *For every $p \leq \mathbf{P}(T > n)$ and for every $q \geq \sum_{i=0}^n \mathbf{P}(T > i)$ we have*

$$\mathbf{E}|X_n| \geq \frac{np\lambda(1/p)}{16}, \tag{10}$$

$$\mathbf{E}|X_n| \leq 4\sqrt{n}(H(Y_n) + H(\text{supp } Q_n | Q_n) + 2q(h(n/q) + \log(n+1)))^{1/2}$$

$$\mathbf{E}|X_n| \leq 3\bar{\lambda}(n/q)q + 12\sqrt{n}\sqrt{H(Y_n) + H(\text{supp } Q_n) + \mathbf{E}|Q_n|}, \tag{11}$$

$$H(X_n) \geq \frac{np^2\lambda^2(1/p)}{4096},$$

$$H(X_n) \leq H(Y_n) + H(\text{supp } Q_n | Q_n) + 2q(h(n/q) + \log(n+1)).$$

Proof of Theorem 2. The lower bound follows from (10) and the almost monotonicity of speed (4). The bound (11), our assumptions (1), and (4) yields

$$\mathbf{E}|X_n| \leq cn(p_n\lambda(1/p_n) + \sqrt{p_n}).$$

Now for a generating random walk on any transitive graph $\lambda(n) \geq c\sqrt{n}$, see Lee and Peres (2009). The upper bound follows. \square

5 The assembly line

The goal of this section is to describe and analyze a simple birth-and-death Markov chain, called the **assembly line**, which is the core of our construction. In the next section, we will show that it essentially is a Schreier graph of a group. However, the first part of the analysis does not need this fact.

Fix a bounded sequence of integers $m_i \geq 2$, and consider the set S of left-infinite words $\dots w_3 w_2 w_1$, where $w_i \in \{0, \dots, m_i - 1\}$, and w_i are 0 for all i large enough. We consider the following random walk on this set. In each step, toss a fair coin; if heads, randomize w_1 ; if tails, randomize the letter w_{i+1} after the first nonzero letter w_i .

This can be thought of as an infinite assembly line, where w_i is the state of the worker at position i , either dozing off ($w_i = 0$) or doing one of the $m_i - 1$ possible jobs. The foreman, standing in front of the line, has the option to nudge the first worker, or shout at the first awake worker in the line to nudge the worker after them. Nudged workers just change their state to one chosen uniformly at random.

This is a reversible random walk on S , and its graph (as we will see) is the Schreier graph of the action of a group denoted \mathcal{M}_m . In order to analyze the speed in the permutation wreath product $\Lambda \wr_S \mathcal{M}_m$, we just have to understand the first return time T of this walk started at $o = \dots 000$.

Fortunately, this walk projects to a birth-and-death Markov chain. Indeed, consider the projection from S to the set of finite binary strings

$$\dots w_3 w_2 w_1 \mapsto \dots \bar{w}_3 \bar{w}_2 \bar{w}_1, \quad \bar{w}_i = \mathbf{1}(w_i > 0).$$

The projection of the walk has the same description as above, except that “randomize” for the \bar{w}_i means

$$\bar{w}_i := \begin{cases} 0 & \text{with probability } 1/m_i \\ 1 & \text{with probability } 1 - 1/m_i. \end{cases}$$

Note also that from each $\bar{w} \in \bar{S}$, there are only two possible new positions one can move to (except o , where there is only one). This implies that the graph structure of this walk is that of the nearest-neighbor graph of natural numbers. Indeed, the structure is a standard Gray code – an ordering of all binary strings so that consecutive strings differ by only one bit.

It is easy to see that the position of a given string in this order is

$$\text{position}(\dots b_3 b_2 b_1) = \dots \overleftarrow{b_3} \overleftarrow{b_2} \overleftarrow{b_1}, \quad \text{where } \overleftarrow{b_i} = b_i + b_{i+1} + \dots \pmod{2},$$

where the last sum has only finitely many nonzero terms.

Thinking of the walk on S as a random walk on a weighted graph, the uniform stationary measure on S projects to a stationary measure on \bar{S} . More precisely,

$$\pi(\dots b_3 b_2 b_1) = \prod_{i=1}^{\infty} (m_i - 1)^{b_i}$$

and the edge weights are within constants of the vertex weights. This automatically gives that the resistance between o and 2^n in \bar{S} satisfies

$$r_\ell \leq \mathfrak{R}(o, 2^\ell) \leq 2m_* r_\ell, \tag{12}$$

where r_ℓ is the sum of the inverse vertex weights for vertices between these two positions, and m_* is the maximum degree. That is,

$$r_\ell = \sum_{b_1, \dots, b_n} \pi^{-1}(b_n \dots b_1) = \sum_{b_1, \dots, b_n} \prod_{i=1}^{\ell} \left(\frac{1}{m_i - 1} \right)^{b_i} = \prod_{i=1}^{\ell} \left(\frac{1}{m_i - 1} + 1 \right)$$

since the sum factorizes. In short,

$$\mathfrak{R}(o, 2^\ell) \asymp r_\ell := \prod_{i=1}^{\ell} \frac{m_i}{m_i - 1}$$

Define $v_\ell = m_1 \cdots m_\ell$. Note that the ball of radius 2^ℓ in S about o contains exactly v_ℓ vertices.

Proposition 15. *For $\ell > 0$, we have*

$$\sum_{i=0}^{v_\ell r_\ell} \mathbf{P}(T > i) \leq 2v_\ell.$$

Proof. We truncate the graph at 2^ℓ , and consider the random walk on the truncated graph together with the random walk on the original graph. These are naturally coupled until the

hitting time τ of 2^ℓ . Let T' denote the first return time in the truncated graph. Then we have

$$T > i \Rightarrow T \wedge \tau > i \text{ or } T > \tau \Rightarrow T' > i \text{ or } T > \tau.$$

So we have (by bounding the $i = 0$ case by 1)

$$\sum_{i=0}^n \mathbf{P}(T > i) \leq 1 + \sum_{i=1}^{\infty} \mathbf{P}(T' > i) + n\mathbf{P}(T > \tau).$$

The sum of the first quantities equals $\mathbf{E}T' = v_\ell$ (since in any connected weighted graph the expected return time to a vertex equals the total volume over the weight of the vertex). Also $\mathbf{P}(T > \tau) = 1/\mathfrak{R}(0, 2^n) \leq 1/r_\ell$. The claim follows if we set $n = v_\ell r_\ell$. \square

Lemma 16. *Let X be nonnegative interger-valued random variable with expectation a satisfying $\mathbf{E}[X - x | X \geq x] \leq a$ for all integers $x \geq 0$. Then*

$$\mathbf{P}(X > a/4) \geq 1/31.$$

Proof. By Markov's inequality for all integers t we have

$$P[X \geq t + ea | T > t] \leq 1/e$$

This implies that for all integers k

$$P[X \geq kea] \leq e^{-k}$$

and so for all positive real numbers k

$$P[X \geq kea] \leq e^{-k+1}$$

This gives

$$\mathbf{E}X - \mathbf{E}[X \wedge ra] = \int_{ra}^{\infty} P[X \geq x] dx \leq e \int_{ra}^{\infty} e^{-x/(ae)} dx = ae^2 e^{-r/e}$$

With $r = 8$ we get

$$\mathbf{E}X - \mathbf{E}[X \wedge 8a] \leq \mathbf{E}X/2$$

and therefore

$$\mathbf{E}[X \wedge 8a] \geq \mathbf{E}X/2$$

Applying Markov's inequality to $8a - X \wedge 8a$ we get

$$\mathbf{P}(8a - X \wedge 8a \geq ka) \leq \frac{7.5a}{ka} = 7.5/k$$

setting $k = 7.75$ we get

$$\mathbf{P}(X > a/4) \geq 1 - \frac{7.5}{7.75} = 1/31.$$

\square

Lemma 17. *Consider the hitting time T' of o of the random walk on the graph truncated at 2^ℓ . We have*

$$\mathbf{E}_{2^\ell} T' \geq r_{\ell-1} v_{\ell-1} (m_\ell - 1).$$

Proof. This is a birth-and-death chain, one of the simplest kind of Markov chains, and the expected hitting times can be computed explicitly. By Levin, Peres and Wilmer (2009) (page 27, formula (2.16)), we have

$$\mathbf{E}_{2^\ell} T' = \sum_{i=1}^{2^\ell} \frac{1}{p(i-1, i) \pi_i} \sum_{j=i+1}^{2^\ell} \pi_j \geq \sum_{i=1}^{2^\ell-1} \frac{1}{\pi_i} \sum_{j=2^{\ell-1}+1}^{2^\ell} \pi_j = r_{\ell-1} v_{\ell-1} (m_\ell - 1).$$

□

Proposition 18.

$$\mathbf{P}(T > r_{\ell-1} v_{\ell-1} / 4) \geq \frac{1}{62 m_* r_\ell}.$$

Proof. We have

$$\text{walk hits } 2^\ell \text{ before } o \text{ and } T^* > t \Rightarrow T > t, \quad (13)$$

where T^* is the time to hit o after the first time of hitting 2^ℓ . The probability of the first event is given by $1/\mathfrak{R}(0, 2^\ell) \geq 1/(r_\ell 2 m_*)$ by (12). We now consider $T^* > t$. Note that T^* stochastically dominates the hitting time T' from 2^ℓ to o in the graph truncated at 2^ℓ . This is because we $T^* = T' + T''$ where T'' is the number of steps outside $[0, 2^\ell]$ before hitting 0. Note that T' also stochastically dominates the hitting times of o in the truncated graph when starting from lower vertices. Thus Lemma 16 gives

$$\mathbf{P}(T' > \mathbf{E} T' / 4) \geq 1/31.$$

Using the lower bound on $\mathbf{E} T'$ from Lemma 17 we get for $t = r_{\ell-1} v_{\ell-1} / 4$ we have $\mathbf{P}(T' > t) \geq 1/31$. By the strong Markov property and (13) we get

$$\mathbf{P}(T > t) \geq \frac{1}{62 m_* r_\ell}.$$

□

Recall that

$$r_\ell = \prod_{i=1}^{\ell} \frac{m_i}{m_i - 1}, \quad v_\ell = m_1 \cdots m_\ell, \quad n_\ell = r_\ell v_\ell$$

We define the inverse function and an exponent

$$\ell(n) = \min\{\ell : n_\ell \geq n\}, \quad \alpha_n = \frac{\log v_{\ell(n)}}{\log n}$$

Note that

$$\begin{aligned} v_{\ell(n)} &= n^{\alpha_n}, \\ n &\leq n_{\ell(n)} \leq 2m_*n, \\ n^{1-\alpha_n} &\leq r_{\ell(n)} \leq 2m_*n^{1-\alpha_n}. \end{aligned}$$

We express the results of Propositions 15 and 18 in terms of n .

Corollary 19.

$$\mathbf{P}(T > n) \geq \frac{1}{500m_*^2} n^{\alpha_n-1}, \quad (14)$$

$$\sum_{i=0}^n \mathbf{P}(T > i) \leq 2n^{\alpha_n}. \quad (15)$$

The combination implies a matching upper bound on (14) and lower bound on (15) up to a constant factor depending on m_* only.

Proof. Since $r_{\ell'-1}v_{\ell'-1}/4 \geq r_{\ell'-2}v_{\ell'-2}$, we set $\ell' = \ell(n) + 2$ to get

$$\mathbf{P}(T > n) \geq \mathbf{P}(T > n_{\ell(n)}) \geq \mathbf{P}(T > r_{\ell'-1}v_{\ell'-1}/4) \geq \frac{1}{62m_*r_{\ell'}} \geq \frac{1}{500m_*^2} n^{\alpha_n-1}.$$

since

$$r_{\ell'} = r_{\ell(n)+2} \leq 4r_{\ell(n)} \leq 8m_*n^{1-\alpha_n}.$$

For the upper bound, we have

$$\sum_{i=0}^n \mathbf{P}(T > i) \leq \sum_{i=1}^{n_{\ell(n)}} \mathbf{P}(T > i) \leq 2v_{\ell(n)} = 2n^{\alpha_n}. \quad \square$$

We finish this subsection by showing that one can get good control over the sequence $\{\alpha_n\}$ by choosing the appropriate degree sequence $\{m_i\}$.

Lemma 20. *For any $\gamma \in [1/2, 1)$ and any function $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ satisfying $f(1) = 1$ and the log-Lipshitz condition that for all real $a, n \geq 1$*

$$a^{1/2}f(n) \leq f(an) \leq a^\gamma f(n)$$

there is a bounded sequence (m_ℓ) such that for all $n \geq 1$ we have

$$c_\gamma^{-1} \leq \frac{f(n)}{n^{\alpha_n}} \leq 2c_\gamma, \quad \text{with } c_\gamma = 3e^{1/(1-\gamma)}.$$

Proof. Note that

$$\frac{f(n_{\ell+1})}{v_{\ell+1}} = \frac{f(n_\ell)}{v_\ell} y \quad \text{where } y = \frac{f(n_\ell m_{\ell+1}^2 / (m_{\ell+1} - 1))}{f(n_\ell)} \frac{1}{m_{\ell+1}}.$$

By the assumptions on f we have

$$\frac{1}{\sqrt{m_{\ell+1} - 1}} \leq y \leq \left(\frac{m_{\ell+1}^2}{m_{\ell+1} - 1} \right)^\gamma \frac{1}{m_{\ell+1}} \quad (16)$$

Set m_* be the smallest integer $m_{\ell+1}$ so that the upper bound at is not more than 1. This depends on γ only. Now, for $\ell \geq 1$ inductively choose

$$m_{\ell+1} = \begin{cases} 2 & \text{if } f(n_\ell)/v_\ell \leq 1 \\ m_* & \text{if } f(n_\ell)/v_\ell > 1 \end{cases}$$

then (16) and induction give the inequality

$$\frac{1}{\sqrt{m_* - 1}} \leq \frac{f(n_\ell)}{v_\ell} \leq 2^{2^{\gamma-1}},$$

which holds trivially for $\ell = 0$.

Equivalently, by definition of α_n , the inequality

$$\frac{1}{\sqrt{m_* - 1}} \leq \frac{f(n)}{n^{\alpha_n}} \leq 2^{2^{\gamma-1}} < 2,$$

holds at the values $n = n_\ell$.

To get the claim for all integers, recall that $n^{\alpha_n} = v_{l(n)}$ is constant between n_ℓ and $n_{\ell+1}$ while f is increasing and may change by a factor at most m_* between n_ℓ and $n_{\ell+1}$, so for all $n \geq 1$ we get

$$\frac{1}{\sqrt{m_* - 1}} \leq \frac{f(n)}{n^{\alpha_n}} < 2m_*.$$

For the value of c_γ note that the right hand side of (16) is at most one when $m_{\ell+1} \geq e \times e^{1/(1-\gamma)}$, and $\lceil e \times e^{1/(1-\gamma)} \rceil \leq 3e^{1/(1-\gamma)}$. \square

6 The piecewise mother groups

The Markov chain random walk on the “assembly line” introduced in Section 5 is a random walk on a Schreier graph of a certain group that we will now introduce.

Given a sequence $m = \{m_i\}_{i \geq 1}$ of integers $m_i \geq 2$, we define \mathbb{T}_m , the spherically symmetric rooted tree with degree sequence \overline{m} to be the following graph. The vertex set given

by finite sequences $a_k \cdots a_1$ where $a_i \in \{0, \dots, m_i - 1\}$, where the empty sequence o is the root. The edge set consists of all pairs of vertices of the form $\{a_k \cdots a_1, a_{k-1} \cdots a_1\}$.

Let $\text{Aut}(\mathbb{T}_m)$ be the set of rooted automorphisms of \mathbb{T}_m , i.e. automorphisms fixing the root. The groups \mathcal{M}_m will be subgroups of $\text{Aut}(\mathbb{T}_m)$.

Automorphisms $\gamma \in \text{Aut}(\mathbb{T}_m)$ can be written as a product

$$\gamma = \langle \gamma_0, \dots, \gamma_{m_1-1} \rangle \sigma$$

where $\sigma \in \text{Sym}(m_1)$ permutes the subtrees of o and γ_i are automorphism of the subtrees. Thus the γ_i are elements of $\text{Aut}(\mathbb{T}_{\tau m})$ where $\mathbb{T}_{\tau m}$ is the tree with the shifted degree sequence $\tau m := m_2, m_3, \dots$. The natural action of the group $\text{Aut}(\mathbb{T}_m)$ on infinite strings (the boundary of the tree) $\cdots a_3 a_2 a_1$ can be defined recursively by

$$\cdots a_3 a_2 a_1 \cdot \gamma = (\cdots a_3 a_2 \cdot \gamma_{a_1}) a_1 \cdot \sigma$$

More generally, any automorphism can be written as a product of automorphisms γ_v of subtrees of vertices v at level ℓ and a permutation permuting vertices at level ℓ . The γ_v are uniquely determined and are called the **sections** of γ at v .

We will be interested in two special types of automorphisms:

1. **Permutations:** automorphisms of the form $\sigma = \langle id, id, \dots, id \rangle \sigma$ which simply permute the children of the root according to σ , with all other sections trivial.
2. **Propagating actions:** For every $\ell \geq 1$ fix some permutation $\sigma_\ell \in \text{Sym}(m_\ell)$. Then we define recursively

$$a_\ell = \langle a_{\ell+1}, \sigma_{\ell+1}, \dots, \sigma_{\ell+1} \rangle id$$

Consider a finite group H of propagating actions with the property that $H^{-1} = H$ and for uniform random $h \in H$ the number $0 \cdot \sigma_\ell^{(h)}$ is uniformly distributed on $\{0, \dots, m_{\ell-1}\}$. This is a variant of automorphisms considered by Brioussel (2011a)).

A simple example is when H is the set of all propagating actions where σ_ℓ depends only on m_ℓ ; in the rest of the paper we will assume that this is the case; this choice affects only the constants in our bounds.

We define \mathcal{M}_m as the group generated by the propagating actions H and the permutations $\text{Sym}(m)$.

We consider the random walk Y_n on \mathcal{M}_m where the step distribution is the even mixture of uniform measures on $\text{Sym}(m_1)$ and H . By definition of H , the random walk $o.Y_n$ on

infinite strings (with finitely many nonzero letters) corresponds to the assembly line walk introduced in Section 5.

Before going on to prove an upper bound on the entropy of the walk X_n (which will then translate to a bound on the speed of X_n), we must further analyze some structures related to the inverted orbit.

The inverted orbit $o, o.Y_1^{-1}, \dots, o.Y_n^{-1}$ is a sequence of rays in the tree \mathbb{T}_m . Their union is a subtree of \mathbb{T}_m called the **ray tree**. Each ray in the ray tree ends with an infinite sequence of zeros.

Lemma 21. *If a vertex v has exactly one child wv in the ray tree then the section at wv is just a propagating action, and the set of descendants of wv in the ray tree is just the 0-ray from wv .*

Proof. Let $|v|$ denote the distance of v from o in the graph metric. Consider the section at v of the word Y_n , written in alternating form $s_v = \sigma_1 a_1 \sigma_2 a_2 \sigma_3 a_3 \dots \sigma_k$, where $\sigma_i \in \text{Sym}(m_{|v|+1})$ and the a_i are propagating actions in the descendant subtree of v .

Since v has only one child wv in the ray tree, we have

$$w = 0.d_1^{-1} = 0.d_2^{-1}d_1^{-1} = \dots = 0.d_{k-1}^{-1} \dots d_1^{-1}$$

This implies that $0.d_2 = 0.d_3 = 0.d_{k-1} = 0$, so the section at wv is just $a_1 a_2 a_3 \dots a_{k-1}$, a propagating action. The set of descendants of v in the ray tree is exactly the union of rays $(0.d_1^{-1})v, (0.d_2^{-1}d_1^{-1})v, \dots, (0.d_{k-1}^{-1} \dots d_1^{-1})v$, i.e. the ray $\dots 0000wv$, as required. \square

We can now show the following claim:

Lemma 22. *The number of ray trees with r infinite rays is at most $(m_* + 1)^{6r}$.*

Proof. Prune the ray tree at each ray after the first lone child. In other words, if v is the first vertex in a ray to have only one child u in the ray tree, then u is in the pruned tree but its descendants are not. Note that the pruned tree determines the ray tree: the ray tree is given by adding 0-rays to the leaves of the pruned tree.

The pruned tree is finite; let i be the number of its vertices. Since we pruned each ray at the first lone child, any vertex in the pruned tree save for the leaves and their parents has at least two children. If the total number of leaves is r , then the total number of children is at least $r + (i - 2r) * 2$, we get the inequality $i \leq 3r - 1$.

The tree can be described by a path that visits every vertex and passes through every edge at most twice; such a path can be described by $2i$ numbers between 0 and m_* giving a list of directions which neighbor to go to next. The number of such lists is at most $(m_* + 1)^{6r}$. \square

Lemma 23. *The number of group elements for which there are words with a given ray tree with r rays is at most $(m_*!)^{3m_*^2 r}$.*

Proof. For a group element g its section at a vertex determines all the sections on the descendant of the vertex. Define the reduced section at a vertex v as the permutation part of the section a in its decomposition $a = \langle a_i \rangle \sigma$.

A finite rooted subtree of \mathbb{T}_m is called full if each vertex either has maximal number of children or no child.

Given a rooted full subtree of \mathbb{T}_m , the sections of g at the leaves together with the reduced sections at the rest of the subtree determine the action of g on \mathbb{T}_m .

Now given a word, consider the pruned version T' of its ray tree. Let T'' be the smallest full subtree of \mathbb{T}_m containing T' . This is called the minimal tree in Brioussel (2011a).

By Lemma 21 the sections at the leaves of T'' are either permutations or propagating actions. The number of vertices of T'' is at most m_* times the number of vertices in T' , which is at most $3r$ by Lemma 22.

The total number of permutations together with propagating actions is bounded above by $m_*! + 2!3! \cdots m_*! \leq (m_*!)^{m_*}$. So the total number of elements for which there is a word with a given ray tree is bounded above by $(m_*!)^{3m_*^2 r}$. \square

The next proposition contains an upper bound on the entropy of the walk on \mathcal{M}_m . The same bound with a multiplicative error term of $n^{o(1)}$ was first proved by Brioussel (2011a) using techniques from Bartholdi et al. (2010).

We take a different approach, describing an element Y_n through the inverted orbit of the random walk word and the sections of Y_n .

This is similar to the approach in Brioussel (2011a), where a slightly different tree is used to describe a word. The ray-tree description of the inverted orbit together with our tight estimates on the expected size of the inverted orbit allow us to get a tighter bound.

Corollary 24. *The conditional entropy satisfies*

$$\begin{aligned} H(\text{supp } Q_n \mid |Q_n|) &\leq 6 \log(m_* + 1) \mathbf{E}|Q_n|, \\ H(\text{supp } Q_n) &\leq 6 \log(m_* + 1) \mathbf{E}|Q_n| + \log(n + 1), \\ H(Y_n \mid |Q_n|) &\leq 5 \mathbf{E}|Q_n| m_*^3 \log(m_*), \\ H(Y_n) &\leq 5 m_*^3 \log(m_*) \mathbf{E}|Q_n| + \log(n + 1). \end{aligned}$$

Proof. The conditional entropy is the expected entropy of the conditional distribution. However, the ray-tree is determines the support of Q_n , and it has $|Q_n|$ infinite rays. Thus

the conditional distribution of $\text{supp } |Q_n|$ given $|Q_n|$ is supported on a set of size at most $(m_* + 1)^{6|Q_n|}$. The entropy of the conditional distribution is bounded by the logarithm of this. The first claim follows by taking expectations, and the second follows by the trivial upper bound $H(|Q_n|) \leq \log(n + 1)$.

For the second claim, note that the number of group elements with a word with a given ray tree with r leaves is at most $(m_*!)^{3m_*^2 r}$ by Lemma 23, so multiplying by the number of r -trees and taking logarithms gives the bound. We used the inequality $6 \log(m_* + 1) + 3m_*^2 \log(m_*!) \leq 5m_*^3 \log m_*$.

For the last inequality, use the conditional entropy formula and the fact that $1 \leq |Q_n| \leq n + 1$, so $H(|Q_n|) \leq \log(n + 1)$. \square

Theorem 25.

$$\begin{aligned} \mathbf{E}|X_n| &\geq \frac{1}{8000m_*^2} n^{\alpha_n} \underline{\Delta} \left(\frac{n^{1-\alpha_n}}{500m_*^2} \right), \\ \mathbf{E}|X_n| &\leq 6n^{\alpha_n} \bar{\lambda}(n^{1-\alpha_n}) + 48m_*^2 n^{\frac{1+\alpha_n}{2}}. \end{aligned}$$

By Lee and Peres (2009), for an infinite group, $\bar{\lambda}(n)$ is at least a $c\sqrt{n}$. In particular, in the second inequality, the first term is dominant. Thus our bounds are tight up to constant factor, since by (4) we have $\bar{\lambda} \leq c'\underline{\Delta}$.

Proof. Corollary 5 gives

$$\mathbf{E}|Q_n| = \sum_{i=0}^n \mathbf{P}(T > i)$$

Corollary 19 gives

$$\mathbf{P}(T > n) \geq \frac{1}{500m_*^2} n^{\alpha_n-1}, \quad \sum_{i=0}^n \mathbf{P}(T > i) \leq 2n^{\alpha_n}.$$

Corollary 24 gives

$$\begin{aligned} H(\text{supp } Q_n) &\leq 6 \log(m_* + 1) \mathbf{E}|Q_n| + \log(n + 1), \\ H(Y_n) &\leq 5m_*^3 \log(m_*) \mathbf{E}|Q_n| + \log(n + 1). \end{aligned}$$

and finally, Corollary 14 gives that for any $p \leq \mathbf{P}(T > n)$ and any $q \geq \mathbf{E}|Q_n|$ we have

$$\begin{aligned} \mathbf{E}|X_n| &\geq \frac{np\underline{\Delta}(1/p)}{16}, \\ \mathbf{E}|X_n| &\leq 3\bar{\lambda}(n/q) q + 12\sqrt{n} \sqrt{H(Y_n) + H(\text{supp } Q_n) + \mathbf{E}|Q_n|}. \end{aligned}$$

Combining these inequalities yields the Theorem. \square

Theorem 26.

$$\begin{aligned} H(X_n) &\geq \frac{n^{2\alpha_n-1} \underline{\lambda}^2(n^{1-\alpha_n})}{2^{30} m_*^4}, \\ H(X_n) &\leq (15m_*^4 + 2h(n^{1-\alpha_n})) n^{\alpha_n}. \end{aligned}$$

Denote h_1 the entropy of the step distribution. Another lower bound, useful for finite or small lamp groups is

$$H(X_n) \geq \frac{h_1}{500m_*^2} n^{\alpha_n}.$$

Proof. By Corollary 14 for every $p \leq \mathbf{P}(T > n)$ and for every $q \geq \sum_{i=0}^n \mathbf{P}(T > i)$ we have

$$\begin{aligned} H(X_n) &\geq \frac{np^2 \underline{\lambda}^2(1/p)}{4096}, \\ H(X_n) &\leq H(Y_n) + H(\text{supp } Q_n | |Q_n|) + 2q(h(n/q) + \log(n+1)). \end{aligned}$$

The first two inequalities follow just as in Theorem 25. For the last inequality, note that for $X_n = (\mathcal{L}_n, Y_n)$ we have

$$H(X_n) \geq H(\mathcal{L}_n | Q_n) \geq h_1 \mathbf{E} Q_n$$

where we used the monotonicity of entropy. \square

Proof of Theorems 1 and 3. Theorem 3 follows directly from Lemma 20 and Theorem 26 for finite lamp groups, since $h(\cdot)$ in this case is bounded by a constant.

For Theorem 1, we could use Theorem 2, but in this case there are more direct arguments. Note that given such f , $\tilde{f}(n) = f^2(n)/n$ satisfies the conditions of Lemma 20. The theorem now follows from Theorem 25, and the fact that simple random walk on \mathbb{Z} satisfies

$$\frac{2}{\pi} \sqrt{k} \leq \underline{\lambda}(k) = \bar{\lambda}(k) \leq \sqrt{k}. \quad \square$$

Remark 27. By following the constants throughout the theorems, we get the following explicit constants for the bounds speed and entropy of the groups constructed in the main theorems. Recalling that in lemma 20 one may take $m_* \leq C_\gamma = 3e^{\frac{1}{1-\gamma}}$.

For Theorem 3 we get

$$\frac{\log 2}{1000 C_\gamma^3} f(n) \leq H(X_n) \leq 16 C_\gamma^{9/2} f(n)$$

and for Theorem 1:

$$\frac{1}{2^{19} C_{2\gamma-1}^{7/2}} f(n) \leq E|X_n| \leq 49 C_{2\gamma-1}^{9/4} f(n)$$

The constants can be improved by a more careful calculation. There are also other points in the proof where more careful choices would improve the constants, such as choosing a different set of generators for the base group and noticing the groups constructed actually have a degree sequence consisting of 2 and m_* only.

7 Open questions

First, it would be nice to get the missing exponents; let us conjecture a positive answer to the question of Vershik and Peres:

Conjecture 28. *For every $\beta \in (1/2, 3/4)$ there exists a random walk on a finitely generated group with*

$$\mathbf{E}|X_n| \asymp n^\beta.$$

It is also interesting what exponents one can get for finitely presented groups, of which there are only countably many. We believe that our examples can be modified to get a dense set of exponents in $[3/4, 1]$, but not to make the exponents oscillate. In this more restricted world, the following is still open.

Question 29. *Let X_n be a random walk on a finitely presented group. Is there necessarily a β so that*

$$\mathbf{E}|X_n| = n^{\beta+o(1)}?$$

Related to this, we also have

Conjecture 30. *The set of β so that there exists a finitely presented group and a random walk on the group so that $\mathbf{E}|X_n| = n^{\beta+o(1)}$ is dense in $[1/2, 1]$.*

The **Hilbert compression exponent** $\alpha^*(G)$ of a group of G is the supremum over all $\alpha \geq 0$ such that there exists a 1-Lipschitz mapping $f : G \rightarrow L_2$ and a constant $c > 0$ such that for all $x, y \in G$ we have $\|f(x) - f(y)\|_2 \geq c \text{dist}(x, y)^\alpha$. Austin, Naor and Peres (2009) showed that the speed exponent $\beta(G)$, if exists, satisfies

$$\alpha(G)\beta(G) \leq 2. \tag{17}$$

Thus every embedding gives an upper bound on the speed exponent. We do not know if this bound can be sharp for our examples \mathcal{M}_m in the case when β exists.

Question 31. *Assume that m is a constant sequence. Is it true that $\alpha(\mathbb{Z} \wr_S \mathcal{M}_m)\beta(\mathbb{Z} \wr_S \mathbb{M}_m) = 2$?*

This would give an alternative upper bound, and apart from an $n^{o(1)}$ fraction of error, would strengthen our theorem. It would also give new examples for Hilbert compression exponents.

The groups we are considering are permutation wreath products by piecewise automaton groups. It is still an open question to give the precise speed, or even entropy exponent on any of the classical automaton groups (except perhaps the lamplighter group); we will skip the definition of these groups but they can be found in, for example Amir et al. (2009).

Problem 32. *Specify the speed exponent or the entropy exponent in any of the following groups: Grigorchuk’s group, Basilica group, Hanoi towers group, degree-zero mother groups.*

Acknowledgments. We thank Omer Angel for useful discussions, as well as Gady Kozma for the idea of Lemma 12 which allowed us to remove a logarithmic error term from our bounds. We also think Péter Mester and Andrew Stewart for carefully reading previous versions. The work of the first author was supported by the Israeli Science Foundation. The work of the second author was supported by the Canada Research Chair program and the NSERC Discovery Accelerator program.

References

- Amir, G., Angel, O. and Virág, B. (2009). Amenability of linear-activity automaton groups, *To appear in the Journal of the European Mathematical Society. Arxiv 0905.2007*.
- Amir, G. and Virág, B. (2011). Positive speed for high-degree automaton groups, *Arxiv 1102.4979*.
- Austin, T., Naor, A. and Peres, Y. (2009). The wreath product of \mathbb{Z} with \mathbb{Z} has Hilbert compression exponent $\frac{2}{3}$, *Proc. Amer. Math. Soc.* **137**(1): 85–90.
- Bartholdi, L. and Erschler, A. G. (2010). Growth of permutational extensions, *Arxiv 1011.5266*.
- Bartholdi, L., Kaimanovich, V. A. and Nekrashevych, V. V. (2010). On amenability of automata groups, *Duke Math. J.* **154**(3): 575–598.
- Bartholdi, L. and Virág, B. (2005). Amenability via random walks, *Duke Math. J.* **130**(1): 39–56.
- Brieussel, J. (2011a). Behaviors of entropy on finitely generated groups, *Arxiv 1110.5099*.
- Brieussel, J. (2011b). Growth behaviours in the range $e^{(r^\alpha)}$, *Arxiv 1107.1632*.

- Carne, T. K. (1985). A transmutation formula for Markov chains, *Bull. Sci. Math. (2)* **109**(4): 399–405.
- Erschler, A. (1999). An example of the rate of departure to infinity for a random walk on a group., *translation in Russian Math. Surveys* **54**(5): 10231024.
- Erschler, A. (2003). On drift and entropy growth for random walks on groups., *Ann. Probab.* **31**(3): 11931204.
- Erschler, A. (2006). Piecewise automatic groups, *Duke Math. J.* **134**(3): 591–613.
- Grigorchuk, R. I. (1984). Degrees of growth of finitely generated groups and the theory of invariant means, *Izv. Akad. Nauk SSSR Ser. Mat.* **48**(5): 939–985.
- Hebisch, W. and Saloff-Coste, L. (1993). Gaussian estimates for Markov chains and random walks on groups, *Ann. Probab.* **21**(2): 673–709.
- Kaimanovich, V. A. and Vershik, A. M. (1983). Random walks on discrete groups: boundary and entropy, *Ann. Probab.* **11**(3): 457–490.
- Kassabov, M. and Pak, I. (2011). Groups of Oscillating Intermediate Growth, *Arxiv 1108.0262*.
- Kesten, H. (1959). Symmetric random walks on groups, *Trans. Amer. Math. Soc.* **92**: 336–354.
- Lee, J. and Peres, Y. (2009). Harmonic maps on amenable groups and a diffusive lower bound for random walks, *Arxiv 0911.0274*.
- Levin, D. A., Peres, Y. and Wilmer, E. L. (2009). *Markov chains and mixing times*, American Mathematical Society, Providence, RI. With a chapter by James G. Propp and David B. Wilson.
- Naor, A. and Peres, Y. (2008). Embeddings of discrete groups and the speed of random walks, *Int. Math. Res. Not. IMRN* **076**.
- Thompson, R. (2010). The rate of escape of random walks on polycyclic and metabelian groups., *Arxiv 1010.0983*.
- Varopoulos, N. T. (1985). Long range estimates for Markov chains, *Bull. Sci. Math. (2)* **109**(3): 225–252.

Gideon Amir
Department of Mathematics
Bar Ilan University
Ramat Gan, 52900 Israel
`amirgi@macs.biu.ac.il`

Bálint Virág
Departments of Mathematics and Statistics
University of Toronto
Toronto ON M5S 2E4, Canada
`balint@math.toronto.edu`